Dynamic Optically Multiplexed Imaging

Yaron Rachlin, Vinay Shah, R. Hamilton Shepard, and Tina Shih
Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street, Lexington, MA, 02420

ABSTRACT

Optically multiplexed imagers overcome the tradeoff between field of view and resolution by superimposing images from multiple fields of view onto a single focal plane. In this paper, we consider the implications of independently shifting each field of view at a rate exceeding the frame rate of the focal plane array and with a precision that can exceed the pixel pitch. A sequence of shifts enables the reconstruction of the underlying scene, with the number of frames required growing inversely with the number of multiplexed images. As a result, measurements from a sufficiently fast sampling sensor can be processed to yield a low distortion image with more pixels than the original focal plane array, a wider field of view than the original optical design, and an aspect ratio different than the original lens. This technique can also enable the collection of low-distortion, wide field of view videos. A sequence of sub-pixel spatial shifts extends this capability to allow the recovery of a wide field of view scene at sub-pixel resolution. To realize this sensor concept, a novel and compact divided aperture multiplexed sensor, capable of rapidly and precisely shifting its fields of view, was prototyped. Using this sensor, we recover twenty-four megapixel images from a four megapixel focal plane and show the feasibility of simultaneous de-multiplexing and super-resolution.

Keywords: optical multiplexing, computational imaging, compressive sensing, super-resolution

1. INTRODUCTION

In traditional imaging systems, an increased field of view comes at the expense of coarser sampling of the scene. Each pixel on the focal plane maps to an area in object space and as such, the field of view can only be expanded by increasing the sampling area. Methods have been developed to create wide FOV systems without sacrificing fine image detail by stitching together images from multiple narrow field of view sensors, by scanning a single narrow field of view sensor across the scene, and through super-resolution techniques that combine a series of images with sub-pixel shifts. Like the vast majority of optical sensors used today, these systems are based on the principle that at any instant in time each pixel views only a single point in object space. In these established approaches, increasing the spatial resolution requires either more pixels or more time.

Optically Multiplexed Imaging is based on the principle that a single pixel can be used to observe multiple object points simultaneously. The image formed by an optically multiplexed system is the superposition of multiple images formed by discrete imaging channels. This has been investigated in designs that use multiple lenses to form images on a single focal plane array (FPA)[1], a cascade of beam splitting elements to divert multiple fields of view into a single lens[2,3,4,5], and by placing an interleaved array of sub-aperture micro-prisms in front of a single lens[6,7]. In a recent paper[8] we presented a new optical design architecture based on a division of aperture technique to divide the pupil area of a single lens into a number of independent imaging channels. This method offers advantages over prior approaches through its flexibility to individually direct and encode the optical channels and it yields a significant volume advantage in systems with a high degree of multiplexing.

A single multiplexed image is inherently compressed, and without additional encoding, contains inherent ambiguities. While it is feasible to detect objects in a multiplexed image, the angular position of a detected object is uncertain since it could have appeared in any of the N image channels. Static encoding schemes, such as changing the point spread function of each channel[8], combining measurements from two or more multiplexing sensors with different multiplexing channel parameters[9], or relying on differential channel overlap and rotation, enable disambiguation and tracking of localized objects[3].

We have developed a method of dynamic encoding that allows for time-varying spatial and/or temporal encoding of the signals from each channel. A dynamic encoding architecture that is both rapid and precise provides a powerful flexibility to optimize a multiplexed imager for a variety of sensing tasks. In many envisioned scenarios the disambiguation task need only be performed intermittently, and so the ability to dynamically activate and deactivate the encoding function allows for optimization of image disambiguation, signal to noise ratio, or frame rate. In addition, dynamic encoding allows for optimization of either sparse scene or dense scene imaging modes. Spatial encoding can be implemented in a high frame rate sparse-scene detection scenario by rapidly shifting channel images to deterministically encode their point spread functions via motion blur. Reconstruction of arbitrary information rich dense-scenes may be accomplished through temporal encoding by precisely shifting and stabilizing independent channel images in a sequence of measured frames. Channel image shifts can be optimized for an undetermined set of measurements that are used in conjunction with assumptions of the scene content for compressive reconstruction. Alternatively, channel image shifts can be optimized for suppression of image artifacts related to motion or noise when a larger number of frames are collected to form a fully determined system of equations for image reconstruction. Fully determined image reconstruction has been previously achieved by using shutters to attenuate individual imaging channels[4] or by using a slow moving element to continuously shift a single channel's image between samples for a two channel multiplexed imager[2]. Our method of dynamic encoding by independently shifting channel images also allows for super-resolved image reconstruction by precisely varying the shift magnitude.

In this paper we demonstrate a dynamic, fast, and precise multi-channel shift-based encoding method for an optically multiplexed imager. Section 2 will demonstrate that it is feasible to use a sequence of per channel shifts to construct a full rank measurement matrix. In addition, we will discuss the algorithmic approach used to approximately invert these large measurement matrices. This section will also discuss how precise, sub-pixel shifts can be used to achieve spatial super-resolution. Section 3 will describe a novel aperture division six-channel multiplexed imaging system that demonstrates our concept. The dynamic and precise encoding in this prototype is achieved via fast and accurate piezo stages. This system is used to collect data for the results presented in Section 4. The results section includes a 24 megapixel image reconstruction from a 4 megapixel focal plane, and the first demonstration of simultaneous de-multiplexing and spatial super-resolution.

2. IMAGE ENCODING AND RECOVERY

An imaging process can be expressed as a linear transformation from object to image space with additive noise. This can be represented as

$$\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x} + \mathcal{E}, \tag{2-1}$$

where $\boldsymbol{z} \in \mathbb{R}^{l\,x\,1}$ is the measured image observed on the focal plane, $\boldsymbol{A} \in \mathbb{R}^{l\,x\,m}$ is the imaging transformation matrix, $\boldsymbol{x} \in \mathbb{R}^{m\,x\,1}$ is a discretized $\boldsymbol{m}$-pixel representation of the scene which we desire to reconstruct, and $\mathcal{E} \in \mathbb{R}^{l\,x\,1}$ represents the noise corrupting each pixel measurement. A multiplexing imaging process has a transformation matrix comprised of an encoding, a selection, a downsampling, and a multiplexing operation. Thus, the multiplexing imaging process can be written as

$$\boldsymbol{z} = \boldsymbol{A}_{\boldsymbol{multiplex}}\boldsymbol{A}_{\boldsymbol{downsample}}\boldsymbol{A}_{\boldsymbol{selection}}\boldsymbol{A}_{\boldsymbol{encoding}}\boldsymbol{x} + \mathcal{E}. \tag{2-2}$$

The encoding operation, $\boldsymbol{A}_{\boldsymbol{encoding}} \in \mathbb{R}^{nm\,x\,m}$, produces an encoded version of the underlying scene for each of the $\boldsymbol{n}$ channels. In this paper, the encoding is a 2-dimensional per channel image shift for each of the channels. The shifts are an integer number of pixels in the reconstructed image space. The selection matrix, $\boldsymbol{A}_{\boldsymbol{selection}} \in \mathbb{R}^{f\ln\,x\,nm}$, represents the mapping from the shifted scene coordinates to focal plane coordinates. The downsampling factor, $\boldsymbol{f}$, is the ratio of the area of a focal plane pixel to that of the reconstructed pixel. Downsampling, $\boldsymbol{A}_{\boldsymbol{downsample}} \in \mathbb{R}^{\ln\,x\,f\ln}$, is the re-sampling from a higher resolution reconstructed image to that of the lower resolution focal plane. When $\boldsymbol{f} = 1$, the resolutions are matched and $\boldsymbol{A}_{\boldsymbol{downsample}}$ is simply an identity matrix. Finally, the multiplexing operation, $\boldsymbol{A}_{\boldsymbol{multiplex}} \in \mathbb{R}^{l\,x\,\ln}$, sums over the $\boldsymbol{n}$ channels which are physically superimposed on the l pixel focal plane. In general, solving for $\boldsymbol{x}$ given $\boldsymbol{z}$ is an

ill-posed problem as $\boldsymbol{l} < \boldsymbol{m}$ due to the multiplexing and downsampling operations. However, taking $\boldsymbol{p}$ images of a static scene with different image shifts can be written as

$$\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{A}}\boldsymbol{x} + \tilde{\mathcal{E}}, \tag{2-3}$$

where $\tilde{\boldsymbol{z}} = \begin{bmatrix} \boldsymbol{z_1} \\ \boldsymbol{z_2} \\ \boldsymbol{z_3} \\ \vdots \\ \boldsymbol{z_p} \end{bmatrix} \in \mathbb{R}^{pl\,x\,1}$, $\tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A_1} \\ \boldsymbol{A_2} \\ \boldsymbol{A_3} \\ \vdots \\ \boldsymbol{A_p} \end{bmatrix} \in \mathbb{R}^{pl\,x\,m}$, and $\tilde{\mathcal{E}} = \begin{bmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \mathcal{E}_3 \\ \vdots \\ \mathcal{E}_p \end{bmatrix} \in \mathbb{R}^{pl\,x\,1}$. If $pl \geq m$ then $\tilde{\boldsymbol{A}}$ can be full rank for appropriately chosen shift matrices, which allows for solving for the scene by inverting $\tilde{\boldsymbol{A}}$,

$$\hat{\boldsymbol{x}} = \tilde{\boldsymbol{A}}^{-1}\tilde{\boldsymbol{z}}. \tag{2-4}$$

If $pl < m$ then $\tilde{\boldsymbol{A}}$ is undetermined, but with properly chosen shifts it contains sufficient structure to recover a restricted class of signals if the proper regularization is used. For example, if the signal is sparse in some basis, then the image can be recovered using standard nonlinear estimators used in compressed sensing[10,11].

When $\boldsymbol{f} < 1$, and shifts are chosen to be sub focal-plane pixel in size, the recovered image is at a higher resolution than the focal plane resulting in a super-resolution reconstructed image. To achieve super-resolution in a system whose resolution is limited by pixel size, the reconstructed pixel size should be matched to the diffraction limited spot size and thus the amount of super resolution achievable is limited by the precision to which the shifts can be measured as well as the optics. The dimension of the underlying scene increases as $\boldsymbol{f}$ is decreased, due to the increased resolution of the reconstructed image, and thus additional measurements are required for constructing a fully determined $\tilde{\boldsymbol{A}}$ matrix.

For large images (e.g. a multi-megapixel image) the $\tilde{\boldsymbol{A}}$ matrix can become so large that a direct inverse is impractical, since computing an inverse scales cubically with the number of elements. However, $\tilde{\boldsymbol{A}}$ is inherently sparse since a shift corresponds to a sparse matrix, enabling a reduction in computational cost. By modeling the $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{A}}^{\boldsymbol{T}}$operations (without having to explicitly compute the matrices) we can use an iterative solver such as LSQR[12] that can approximate rather than directly computes the inverse. The results shown below used the MATLAB implementation of the LSQR algorithm developed by Stanford's Systems Optimization Laboratory[1413] with modifications allowing it to run on a GPU.

To provide intuition regarding the above equations, we consider the image shown below in Figure 2.1, to which the $\tilde{\mathrm{A}}$ and $\tilde{\mathrm{A}}^{\mathrm{T}}$operations are applied. The image is 100 pixels high and 300 pixels wide, and we are simulating a 100 by 100 detector array multiplexed three times to cover the full field of view. The application of $\tilde{\mathrm{A}}$ yields three multiplexed images, shown in Figure 2.2. There is a relative shift difference between all three channels of less than ten pixels horizontally in each of the multiplexed frames. The application of $\tilde{\mathrm{A}}^{\mathrm{T}}$ to these multiplexed measurements is shown in Figure 2.3. This single application of the transpose matrix shows the starting point of the LSQR algorithm, and demonstrates that some regions are observed more frequently than others as can be seen by their increased intensity. A single application of the transpose operation does not resolve the ambiguities inherent in multiplexed imaging since the $\tilde{\mathrm{A}}$ matrix is not orthogonal.

Figure 2.1 Wide angle scene imaged in simulation

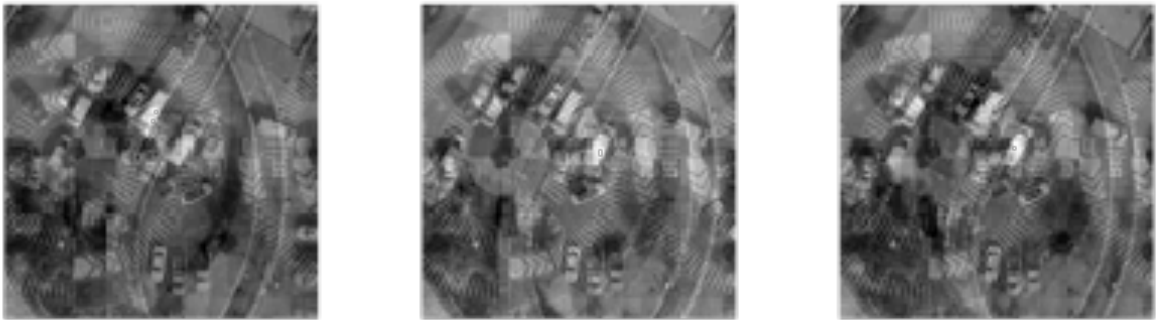


Figure 2.2 Three multiplexed frames corresponding to three multiplexed measurements of the scene, each with its own set of spatial shifts.

Figure 2.3 The result of the transpose of the imaging matrix applied to the measurements

Given the three measurements shown in Figure 2.2, an approximate solution is computed using LSQR as described above, yielding the reconstruction shown in Figure 2.4. The pixel-wise error between the reconstruction and the original image shown in Figure 2.1 is shown in Figure 2.5.

Figure 2.4 Reconstruction of wide-angle scene from multiplexed shift-encoded measurements

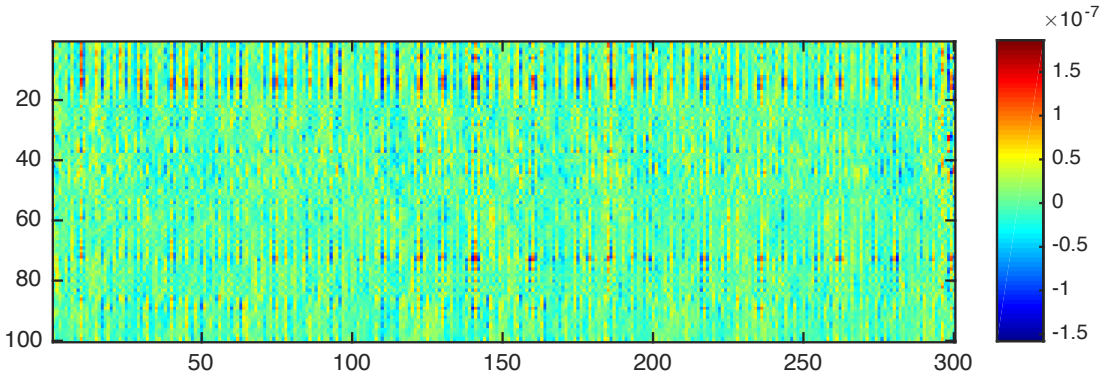


Figure 2.5 Pixel error between the original image and reconstruction.

The uniformly small errors in this reconstruction of a natural image suggest that it is feasible to obtain a full rank measurement system with appropriate image shifts. To confirm this, we simulated 1000 $\tilde{\boldsymbol{A}}$ matrices, where the spatial shifts were chosen randomly. We found that 832 of 1000 were full rank. In this same simulation, we found that key quantities predicting performance under noise, such as matrix condition number, varied widely with shift selection.

Given that independent shifts per channel are sufficient to form full rank measurements matrices, with $\boldsymbol{n}$ multiplexed frames for $\boldsymbol{n}$ multiplexed channels there is no need to make assumptions about the spatial content of a scene. Instead, there is an assumption that the scene is static for a fixed period of time. Therefore, for sufficiently fast encoding and sampling, the imaging capability of a dynamic shift-encoded multiplexed sensor enables imaging of a broad set of scenes. Section 3 below discusses a design that enables such rapid, precise, and independent shift encoding per channel.

3. OPTICAL SYSTEM

An experimental optically multiplexed imaging system was constructed by dividing the entrance pupil of a single large-aperture parent lens into six sub-aperture imaging channels. This aperture division architecture[14] was selected because it provides the most compact and practical method of multiplexing multiple channels onto a single image plane. A Nikon AF-S NIKKOR 200 mm F/2G ED VR II lens was used as the parent lens for this experiment. The imaging camera was a Point Grey Grasshopper3 with a 2048x2048 array of 5.5 micron pixels. Together these produced a 3.2°x3.2° 4.1 Mpix field of view for each channel.

The multiplexing optical element consisted of an array of mirrors that was placed in front of the parent lens. By using planar multiplexing optics in collimated space every channel was ensured to focus to a common image plane. This multiplexing technique allows for the optical system to provide staring coverage over a field of view much wider than the aberration-corrected field of view of the parent lens. Thus, a narrow field of view telephoto lens might operate as a wide field of view panoramic lens. Furthermore, the geometric image distortion in each channel is approximately equal

to that of the parent lens, which offers the potential for extremely wide field of view images without the characteristic distortion of fish-eye lenses.

Each mirror was tilted at a different angle to arrange the multiplexed field of view in a 19°x3.2° panoramic format. The pupil was segmented into six sections as shown in Figure 3.1(b). Each mirror facet was sized to divide the F/4 pupil of the parent lens into 6 equal area sections. The remote location of the mirror assembly with respect to the aperture stop introduced channel-dependant vignetting; however, this effect was minimal due to the small field angles and relatively short distance from the entrance pupil to the multiplexing element. The resulting image irradiance non-uniformities were measured in each channel by flood-illuminating the system with N-1 channels masked. Results were then applied as gain correction maps in the image reconstruction process.

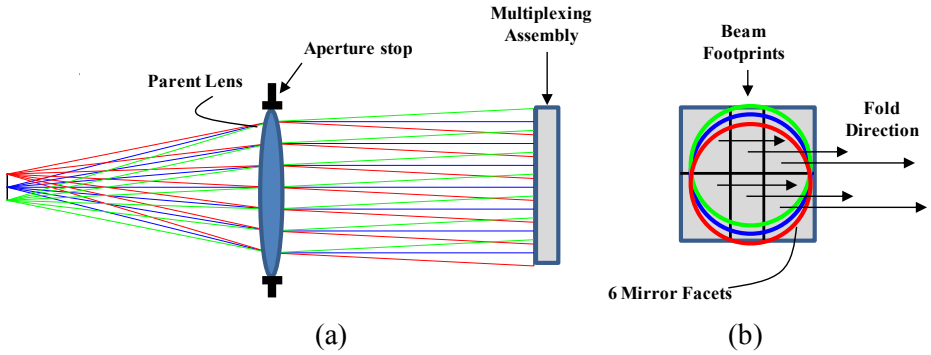


Figure 3.1 Pupil Division Strategy

When dividing the entrance pupil of a parent lens into N equal area channels the effective F/# of each channel scales by sqrt(N). Using the parent lens at an F/4 aperture produced six channels with an effective aperture of F/9.8. The optical resolution in the sagittal and tangential orientations differed because the apertures were non-circular. The effect of the pupil division on MTF is shown in Figure 3.2(a). Parent lens MTF was tested using the ISO 12233 tilted edge method. Results show that the MTF of the full F/4 aperture is greater than 45% at the Nyquist frequency (90.9 lp/mm). This measurement along with an analysis of MTF in the sub-pupil channels indicated that the image resolution in each channel was still detector-limited after the F/# scaling. Therefore, the six-layer multiplexed image can be disambiguated to achieve a full 6x pixel resolution increase with respect to the focal plane's sampling resolution. Further, analysis indicated that positive contrast will remain beyond the Nyquist frequency, which allows for super-resolved image reconstruction.

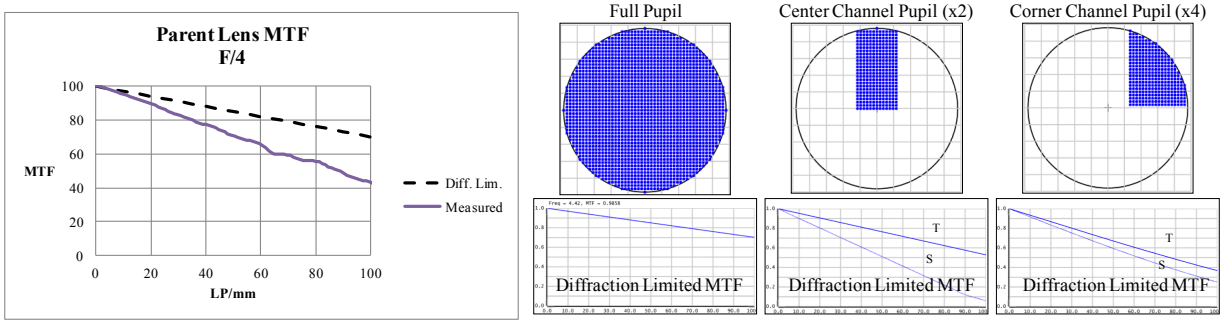


Figure 3.2. Parent lens MTF and pupil sampling. (a) diffraction-limited and measured MTF at F/4. (b) Pupil division and its effect on the diffraction-limited MTF. A difference between the tangential (T) and sagittal (S) MTF is observed in the divided pupils. MTF is plotted out to 100 lp/mm.

Each mirror was mounted on an nPoint RXY3-276 tip/tilt piezoelectric actuator as shown in Figure 3.3. This allowed the mirrors to be steered over a 3 mrad range with 0.05 microradian accuracy. A settling time of 3 miliseconds allowed the mirrors to rapidly step between angles and stabilize with sub-pixel accuracy during the camera readout period. Image blurring due to mirror motion was therefore negligible. Using the actuators, images from individual channels could be independently shifted between frames for dynamic encoding. Image reconstruction of an arbitrary scene was made possible through a sequence of integer pixel-shifts, and sub-pixel shifts allowed for super-resolved image reconstruction.

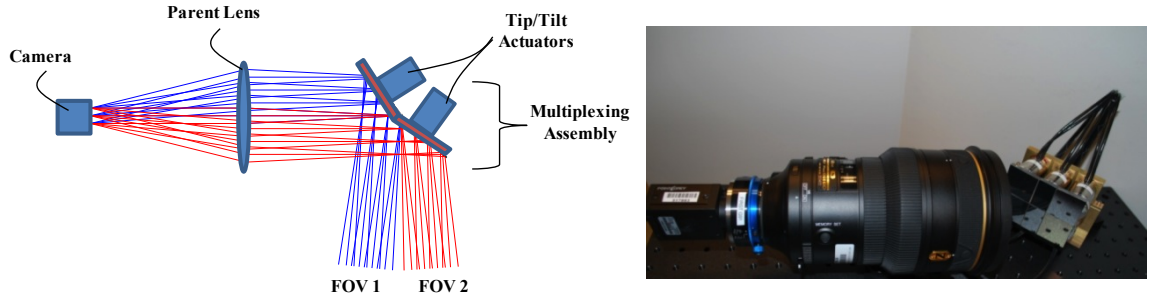


Figure 3.3. Optical System. (a) notional design. (b) prototype system

4. EXPERIMENTAL RESULTS

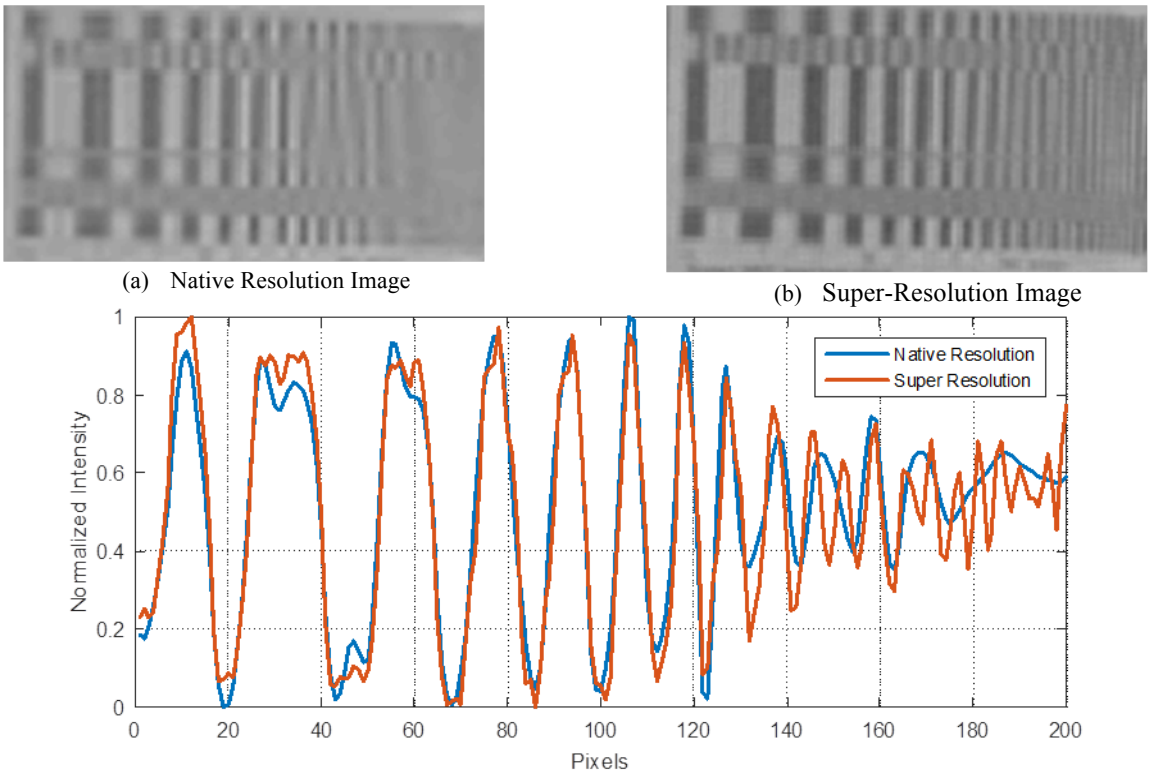
The prototype system described above was used to collect multiple frames of multiplexed image data. Figure 4.1 shows a single 6-channel multiplexed frame. The image resolution is 2048x2048, the native resolution of the camera.

Figure 4.1 Four-Megapixel six-channel multiplexed image

The images were reconstructed to a 24 megapixel image as show in Figure 4.2. Fundamentally, only 6 frames are required for reconstruction, however a collection of 40 frames was used for this reconstruction to reduce error introduced by noise. The need to collect frames beyond the number of channels depends on scene illumination, detector characteristics, and the shift selection. Shown below the reconstructed images are zoomed in regions highlighting that the fine detail in the image is preserved.

Figure 4.2 Reconstructed 24-megapixel scene

The prototype system was also used to demonstrate super-resolution. A resolution bar target with increasing spatial frequency was imaged. A super resolution image was formed with the downsample factor f=0.25 as shown in **Error! Reference source not found.**b. The image was also reconstructed at the native resolution and up-sampled by a factor of two in each dimension for comparison in **Error! Reference source not found.**a. **Error! Reference source not found.**c shows a line out across both images. Greater contrast and spatial frequencies are observed in the super resolution relative to the native resolution image.



(a) Native Resolution Image

(b) Super-Resolution Image

(c) Lineout of Resolution Images

Figure 4.3 Super Resolution Results. (a) Reconstructed image at the camera's native resolution and then 2x interpolated using bicubic interpolation in each dimension. (b) Reconstructed at 2x native resolution in both dimensions. (c) Line out showing increased contrast and resolvability in super resolution reconstruction relative to native reconstruction.

5. CONCLUSION

This paper demonstrates a novel method for imaging via rapid and precise dynamic shifting in a multiplexed sensor. Pairing this technique with a sufficiently fast sampling sensor enables low distortion imaging with more pixels than the original focal plane array, a wider field of view than the original optical design, and an aspect ratio different than the original lens. This technique can also enable the collection of low-distortion, wide field of view videos. A sequence of sub-pixel spatial shifts extends this capability to enable the recovery of a wide field of view scene at sub-pixel resolution. Rapid and precise shifting can be realized via a novel, compact, and practical division of aperture multiplexed sensor. The prototype presented in this paper demonstrated these concepts by recovering twenty-four megapixel images from a four megapixel focal plane, and showing the feasibility of simultaneous de-multiplexing and spatial super-resolution.

REFERENCES

[1] M. D. Stenner, P. Shankar, and M. A. Neifeld, "Wide-Field Feature-Specific Imaging," in Frontiers in Optics, Optical Society of America, (2007).
[2] R. F. Marcia, C. Kim, C. Eldeniz, J. Kim, D. J. Brady, and R. M. Willett, "Superimposed video disambiguation for increased field of view," Opt. Express **16**, 16352-16363 (2008)
[3] S. Uttam, N. A. Goodman, M. A. Neifeld, C. Kim, R. John, J. Kim, and D. Brady, "Optically multiplexed imaging with superposition space tracking," Opt. Express **17**, 1691–1713 (2009).
[4] V. Treeaporn, A. Ashok, and M. A. Neifeld, "Increased field of view through optical multiplexing," Opt. Express **18**, 22432-22445 (2010).
[5] R. Horisaki and J. Tanida, "Multi-channel data acquisition using multiplexed imaging with spatial encoding," Opt. Express **18**, 23041-23053 (2010).
[6] C. Y. Chen, T. T. Yang, and W. S. Sun, "Optics system design applying a micro-prism array of a single lens stereo image pair," Opt. Express **16**, 15495-15505 (2008).
[7] A. Mahalanobis, M. A. Neifeld, V. K. Bhagavatula, T. Haberfelde, and D. Brady, "Off-axis sparse aperture imaging using phase optimization techniques for application in wide-area imaging systems," Appl. Opt. **48**, 5212-5224 (2009).
[8] H. R., Shepard, Y. Rachlin, V. Shah, and T. Shih "Design Architectures for Optically Multiplexed Imaging," in submission
[9] R. Gupta, P. Indyk, E. Price, and Y. Rachlin, "Compressive sensing with local geometric features", Proc. of the 27th annual ACM symposium on computational geometry, 87-98, ACM (2011).
[10] Candes, E.J.; Tao, T., "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?," *Information Theory, IEEE Transactions on* , vol.52, no.12, pp.5406,5425, Dec. 2006
[11] Donoho, D.L., "Compressed sensing," *Information Theory, IEEE Transactions on* , vol.52, no.4, pp.1289,1306, April 2006
[12] C. C. Paige and M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, TOMS 8(1), 43-71 (1982).
[13]"Systems Optimization Laboratory." *LSQR: Sparse Equations and Least Squares*. Web. 1 July 2015. <http://web.stanford.edu/group/SOL/software/lsqr/>.
[14] A. Daniels, "Infrared Systems – Technology & Design", SPIE SC835, 279 (2015)

DISCLAIMER: This work is sponsored by the Department of the Air Force under Air Force Contract #FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government.